VOLUME 55
NUMBER 2

WHOLE NO. 250
1943

Psychological Monographs

JOHN F. DASHIELL, *Editor*

Changes in I.Q. at the Public School Kindergarten Level

By

GELOLO McHUGH, PH.D.
Barnard College

Published by
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Publications Office
NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

ACKNOWLEDGMENTS

PLANS of this research were first formulated in discussions with the late Professor Leta S. Hollingworth.

The author is grateful to his advisor, Professor H. E. Garrett, for continued assistance throughout the course of experimentation. He is indebted also to the other members of his advisory committee: Professor H. L. Hollingworth, Professor Otto Klineberg, Dr. T. A. Jackson and Professor R. S. Woodworth for many constructive suggestions and criticisms.

Dr. Charles Selser, Superintendent of Education for the schools in which this research was done, is thanked for his interest and cooperation throughout the testing periods.

The author also wishes to express appreciation for assistance with many details of this report, to his wife, Elizabeth McHugh, and to his friends, Miss Louise Simmons and Miss Jean Ackermann.

TABLE OF CONTENTS

CHANGES IN I.Q. AT THE PUBLIC SCHOOL KINDERGARTEN LEVEL

CHAPTER I

THE PROBLEM

THIS REPORT presents objective data, gathered under carefully controlled conditions, on the question of changes in intelligence quotients at the preschool level. The 1937 Revision of the Stanford-Binet Test was used.

Information on this question has been sought in four directions, all of which are bound up in the general question: Do significant changes in the measured I.Q. occur as a result of preschool experience? The first problem is to make sure whether such changes do occur and this leads to several specific sub-questions as follows: a) To what extent do such changes occur and in which direction? b) How long a period of preschool experience is required to bring about these changes?

The remaining problems are concerned with the explanation of such changes in I.Q. as do occur as a result of preschool experience. The second problem asks whether the gains in I.Q. are accounted for by improved rapport between examiner and subject in the final test as compared with the initial test. a) Can the examiner judge the rapport, in terms of the subject's manifest shyness and negativism, accurately enough to predict in individual instances the gain in I.Q. to be expected from improved rapport? b) Does the change in rapport, as inferred afterwards from the examiner's recorded judgments of shyness and negativism, account for the change in I.Q.?

The third problem asks whether positive changes in I.Q. at this level may be accounted for in terms of socio-economic background and previous environmental experiences of the subject.

The fourth problem asks whether an analysis of individual items of the test according to the subject's approach to them or their requirements of the subject will explain changes in I.Q.

Chapter II

Background of the Problem

THE IOWA REPORTS

The current controversy over the effect of preschool experience on the intelligence quotient had its beginning with the publication in 1932 of an article by Wellman (40) of the University of Iowa Child Welfare Research Station entitled "Some new bases for interpretation of the I.Q." In this article and in subsequent articles (40 through 51) as well as in articles published in the popular press, Wellman maintains that she has presented ample evidence of substantial gains in intelligence made by children attending the preschool laboratories of the Iowa Child Welfare Research Station. These gains are attributed directly to preschool experiences of the children and not to such factors as practice effect with the test, testing techniques, faulty standardization of the tests at the preschool age levels, low reliability of intelligence tests at these levels, or errors made in the statistical approaches to the data obtained. Evidence is offered (42) to demonstrate that gains are associated with preschool attendance since children of preschool age not attending preschool did not make comparable gains. It is also demonstrated that (40) children do not gain in I.Q. during the summer vacation period. It is further demonstrated (43) that length of preschool experience beyond one year bears a direct relation to the amount of gain in intellectual capacity. In addition to length of preschool attendance the following factors are also listed (48) as effecting changes in I.Q.: 1) The child's relative place in his group; 2) Differences in educational practices and procedures; 3) Association with children of high ability; 4) Freedom from personality difficulties. All of the Iowa reports on change in intelligence due to preschool attendance show that gains in I.Q. are much more often obtained from children testing initially near average (100 I.Q.) or below average, while children with definitely superior I.Q.'s on their initial tests are more likely to show losses, when retested after an interval of preschool experience. In many of her published reports (42, 44, 46, 48, 50, 51) Wellman has offered an explanation for this tendency of low and average I.Q.'s to gain and high I.Q.'s to lose from test to re-test after preschool experience in terms of the stimulus value of the environment in which the children spend the interval between tests. She believes that children of initially high I.Q. did not find in the environment sufficient stimulus to intellectual growth and therefore regressed in intelligence in the direction of the average intelligence of the group. The following quotation (50, p. 36) will more nearly summarize Wellman's position:

> Theoretically if the child's I.Q. is as high as the particular environment is capable of producing in him, there will be no increase; conversely, if the environment is a depressing one, the amount of drop will be greater when his initial status is higher.

Wellman believes that her findings have a considerable number of implications both for education and for the mental testing of children (44, 46) and

that our concept of intelligence is in need of drastic revision. She believes that her findings

> make it clear that children do change in test I.Q. from average to genius and from average to feeble-minded and that it must follow that the hereditary limits within which change is permitted must be set much further apart than the customary 4 to 5 points for the probable error of the I.Q.

She offers instead of the concept of inateness of I.Q. what she calls a functional view of intelligence. According to this view the child's I.Q. at a given age is not necessarily of any value with regard to his future I.Q. or necessarily related to his past I.Q.

> One is not bound by the idea that I.Q. is fixed and unchangeable. It happens that the future I.Q. can be predicted if information regarding the stimulus value of the environment in which the child will grow is also available.

Wellman believes (46) that individual intelligence tests are still the best single indicator of the child's general mental ability at or reasonably near the time the measurement is taken.

Coffey and Wellman (9) have published a report on research with data from Iowa preschool children dealing with the relationship between I.Q. gains and the cultural status of the parents of the subjects. According to this report,

> cultural status as measured by the occupation of the father and educational level of parents, while related to the intelligence of the children on entering preschool, was not a significant factor in changes of intelligence of children while attending preschool.

Since the publication of Wellman's report in 1932 there appears to be only one piece of published research other than the Iowa studies which supports the findings of the Iowa researches. This study (35) by Starkweather and Roberts deals with changes in I.Q. of children attending the preschool of the Merrill-Palmer School in Detroit. The authors state that their aim is to verify the findings of Wellman and in order to do so they have taken over their statistical procedures, etc., directly from the Wellman studies. Their conclusion is that their results indicate general corroboration of the Wellman findings concerning the relation between nursery-school attendance and I.Q. changes.

CRITICISMS OF THE IOWA REPORTS

The publication of the Iowa researches has resulted in a number of direct attacks upon them as well as the publication of a considerable number of investigations in the area of intelligence testing at preschool ages. The principal direct criticisms have been made by Simpson (32, 33, 34), McNemar (27, 28), and Goodenough (14, 15, 16, 19). These writers have criticised the Iowa studies in the following major respects: 1) Failure to present the mental development of the same individual from age to age, and presentation of average I.Q. of rapidly diminishing or changing groups, thus ignoring selective factors operating. 2) No consideration given influence of hereditary factors in accounting for individual differences from group to group. 3) Changing tests without keeping separate records of children given one or the other. 4) Vagueness in reporting the selection of groups under consideration. 5) Errors in tabulation. 6) Failure to make mention of rapport factors which might have operated to depress the initial I.Q. 7) Use of incorrect or inadequate methods to determine statistical significance of obtained differences. 8) No account taken of tendency of re-test measurements to

show regression towards the mean of the population with the result that artifactual changes have been attributed to the effect of the environment. 9) Frequent positive conclusions without foundation in facts reported in the studies. 10) The acceptance of the hypothesis "That intelligence tests given to infants and young children have the same predictive value for later mental development as those given after school age." 11) Overlooking the possibility of systematic error in testing as a result of the examiner's knowledge of previous test results for individual children and for the group.

Although not directly aimed at the Iowa reports, two articles by Anderson (1, 2), have contributed to the general criticism. This author analyzed infant and preschool tests with respect to the danger of concluding from them that environment is very important in determining intelligence at early ages. A general summary of his findings follows. 1) The earlier the measurements are made, the less reliance can be placed on a single measurement or observation—if that measurement or observation is used for predicting subsequent development. 2) The earlier the measurements are made, the greater care should be taken to secure accuracy of observation and record, and to follow standardized procedures. 3) The earlier measurements are made the greater account should be taken of the possibility of disturbing factors such as negativism and refusals that operate as constant errors to reduce scores. 4) Since development is a timed series of relations or sequences, there are for many functions periods below which only a small portion of the function can be measured and above which a progressively larger portion can be measured. Hence, the possibilities of prediction are limited and progression with age is not an infallible indication of the value of a measurement.

The present writer might add to the above criticisms to a small extent by stating that a careful search of the Iowa studies, those of Wellman in particular, has revealed no mention of research done by others on factors affecting the intelligence scores of preschool children, for example, research such as that of Rust (31), published in 1931. In this study the author shows that of a group of three-year-old nursery school children 70% showed resistance to items on the Kuhlmann-Binet and that of this resistant group 7% gained from 25 to 35 points in I.Q., 18% 15 to 24 points, 26% 5-14 points, and 14% 1-4 points on re-test on the items after resistance to them was removed. Caille (7) (1933), in an investigation of the resistant behavior of preschool children towards other people, defined resistance as failure to comply with suggestions or an expression of refusal to a person by means of physical or vocal reactions. This author reports that the number of instances of resistance per five minutes was about seven times as large in intelligence test records as in observation or language records obtained in her study. Levy and Tulchin (26) (1925) define resistance in infants and children as any behavior that interferes with the mental test situation and find that the most characteristic resistant reaction for all ages of children from 6 to 54 months is withdrawal reactions plus passive movements.

REPORTS NOT IN AGREEMENT WITH THE IOWA FINDINGS

Since the publication of Wellman's 1932 article (40) a number of reports or findings not in agreement with her have appeared in the literature. These reports

of Anderson (3), Bayley (4, 5), Bird (6), Frandsen and Barlow (12), Goodenough and Maurer (18), Lamson (25), Olson and Hughes (29), and Voas (39) come from a large portion of the important child research centers throughout the United States.

These authors agree in reporting that they are unable to conclude from the results of their researches that preschool experience has any appreciable effect upon the I.Q.

No attempt will be made to review these studies fully at this time. Specific references will be made to certain of them in the discussion of results of the present report.

Chapter III

The Present Investigation

THE SUBJECTS

THE SUBJECTS of this study were the entire entering public school kindergarten classes of a small town in New Jersey. The town (located about twenty-five miles from New York City) supports four public schools, three elementary schools and one large high school. There are no private kindergartens or nursery schools in the town, so nearly one hundred per cent of children of kindergarten entering age attend the public schools.

The population studied appears to be somewhat selected. There are only two parents out of the whole group of parents of subjects who can be classified as laborers on the Barr Occupational Scale (36, p. 66), and information gathered from school records indicates that a large percentage of parents are home owners.

The town's public schools are like other public schools in towns of comparable size throughout the northeastern United States. As in many other towns the seating capacity of the schools has not kept pace with the growing school population. This is definitely true with regard to the kindergarten classes. There are three classrooms, one in each grammar school, available for kindergarten. These classrooms are average size public school classrooms equipped with small tables and chairs for the pupils and with cupboards around the wall in which are stored the usual kindergarten materials used in the public schools. With the exception of games and activities carried on in the school yard during recess periods all kindergarten activities take place in these classrooms. The town's entering kindergarten group has averaged slightly over one hundred pupils for the past several years. Each class is taught by one teacher with the exception of the occasional assistance of a trainee from one of the state normal schools.

Through the cooperation of parents and with the endorsement of the superintendent of schools, it was possible to secure for testing prior to the opening of schools 97 of the 107 kindergarten children registered for the 1940 fall term. Of these 97 children three refused to be tested at this time but were tested during the first day of school and appeared then to be anxious to cooperate with the examiners. By the end of this study, complete data had been obtained on 91 of the 97 children.

The subjects of this study, then, consist of 91 out of 107 entering public school kindergarten children of a small town located near New York City. The group is homogeneous as to chronological age and about evenly divided between the sexes. The mean C.A. of the group at the time of the first test was 62 months, with a S.D. of 4.38 months. There were 44 boys and 47 girls.

THE MEASURING INSTRUMENTS USED

a) Mental Tests Administered to Subjects

1. Stanford-Binet, Revised 1937, forms L and M. (37).

2. Goodenough, Drawing-a-Man Test (17).

b) Observations Made on Subject by Examiner

1. *Free spontaneous verbalization prior to beginning of the Binet Test.* The

examiner recorded in terms of the following descriptions the amount of voluntary talking done by the subject prior to the beginning of the tests. 1) None; 2) Little; 3) Moderate; 4) Free; 5) Incessant.

2. *Judgments regarding the amount of shyness and negativism manifested by the subject during the test situation.* These judgments were marked off on a five point scale represented by a line 120 mm. long, divided into segments of 30 mm. The five points of the shyness scale are: 1) Extreme Shyness: indicating that the subject showed practically no confidence in the test situation; 2) Moderate Shyness: indicating that the subject showed little confidence in the test situation; 3) Shy and Confident: indicating that shyness and confidence were shown about equally in the test situation; 4) Rarely Shy: indicating that the subject was confident in most of the test situation; 5) Never Shy: indicating that the subject was entirely confident in the test situation.

The five points of the negativism scale are: 1) Performs Few Requests; indicating that the subject refused most requests when first presented, that he showed active objections to the requests of the examiner; 2) Usually Performs Requests: indicating slow response to requests and response with evident unwillingness; 3) Performs Requests: indicating a passive attitude towards examiner's requests, and the subject may be said to be neither willing nor unwilling to perform requests; 4) Performs Requests Willingly: indicating that subject was quick to respond to requests, that he responded with evident willingness; 5) Performs Requests with Enthusiastic Willingness: indicating that the subject performed all requests with evident enthusiasm for the performance.

These scales were printed on the examiner's record sheet for each child, a copy of which will be found at the end of this section.

Shyness and Negativism were further defined by the examiner in the following manner:

> It is understood that the examiner will be judging the child on these scales for shyness and negativism in terms of the child's manifestation of these traits while meeting with a strange adult in a social situation. It is clearly understood that the examiner's judgment may be registered at any distance along a scale. It is not necessary to record judgments at the points of a scale.

c) Other Tests Used

1. *The Vineland Social Maturity Scale, Revised.* (11) The information for this scale was obtained through an interview with the subject's mother.

2. *Barr Occupational Rating Scale* (36 p. 66).

3. *The Whittier Home Rating Scale* (52 and 36 p. 73).

d) Other Information Obtained

1. *The record sheet.* This sheet which is reproduced on the page following was partially filled in at the time of the first tests. The remaining information was obtained at the time of the home visit which was made to establish scores on the Vineland, Barr, and Whittier Scales. All parents were visited by one examiner during the first two weeks of the school year.

THE PROCEDURE

The children were brought to a school by appointment. When a child entered the school with his parent he was met by one of the examiners who, after the usual introductory remarks with the parent and the filling in of a record sheet, in-

Subject's No.
Test Date

RECORD SHEET

Subject's Name: Birthdate:
Address:

Parents	G.S.	H.S.	Education Col.	Grad.	Prof.
Mother:					
Father:					

Father's Occupation:
Mother's Occupation:
Appointment Home Visit:
Brothers, Age nearest birthdate:
Sisters, Age nearest birthdate:
Others living in home relationship:

S	Extreme Shyness	Moderate Shyness	Shy and Confident	Rarely Shy	Never Shy
N	Performs Few Req.	Usually Performs Req.	Performs Req.	Performs Req. Willingly	Performs Req. Enthusiastically

vited him to see the kindergarten room like the one in which he would soon be going to school and "to play some school games." The child was never hurried and if he showed resistance at first the parent was asked to go in with him to see the room but was told at the time of the invitation that he would have to withdraw when the child and examiner were ready "to play school games." The child was first shown around the room and viewed some of the kindergarten materials. If it had been necessary for the parent to go to the room the child was then asked if he would stay in the room "to play school games" with the examiner while his parent waited outside. All children except three were willing to stay alone with an examiner. These three were not tested at this time.

At the time of the initial testing of 91 children on whom complete data were obtained 45 were tested on Form L of the 1937 Revision of the Stanford-Binet and 46 on Form M. Each child's final test was on the form alternate to that of his initial test.

During both testing sessions the examiner recorded before administering the first item of the Binet Test the amount of spontaneous verbalization by the child up to that time in accordance with the pre-arranged method described under Measuring Instruments Used.

At the completion of the Binet Test the examiner gave the Goodenough Drawing-a-Man Test (17) twice in succession, the child being provided with a fresh sheet of paper for the second test, the initial effort being removed from sight. The data from these tests are to be treated in a separate publication. During the performance of these last tests the examiner recorded on the scales for shyness and negativism, at the bottom of the record sheet, his judgment of the degree to which the child manifested these traits during the testing procedure.

All tests were administered to 43 of the children by the author. Forty-eight children were tested by an assistant. Every child was tested initially and finally by the same examiner. Data, tests and ratings secured by home visits were ob-

tained by one examiner for the ninety-one cases.

An assistant was secured to test approximately half of the children for two reasons: a) in order that all initial tests might be given within a shorter period before the beginning of school, b) in order that the author might check on the possibility that, even without knowledge of results until final tests were completed, he might unconsciously prejudice results in the direction of his hypotheses regarding changes in I.Q. at the preschool level. Because of this second reason it was decided to secure an assistant to administer all intelligence tests to half of the subjects under the direction of the author without information as to the aims of the research.

In planning this procedure the author was in agreement with Goodenough's statements (18, p. 167) that:

> There have been many studies of the so-called "halo effect" upon ratings, but the possibility of a similar halo resulting from knowledge of a child's standing on a previous test or other facts that predispose the examiner to expect him to do well or to do poorly has been little investigated. Yet any fair-minded and experienced examiner will agree that this factor may be a very real source of error in testing unless definite attempts are made to bring it under control.

Scoring of Tests

One hundred record sheets were prepared for each examiner. These sheets were numbered from 1 to 100 and placed in random order in a binder. The children's names, addresses, C.A.'s, sex, date of the test, and other information were recorded on these sheets at the time of each test. The number of the sheet only was recorded on the test blank. Examiners made full records of children's responses to test items but no tests were scored until all initial and final tests had been given. When all tests were completed each examiner's numbered blanks were taken from a file and carefully shuffled before any scoring was done. Two controls were made possible by this procedure: a) A scorer did not know which child's test was being scored; b) A scorer did not know whether he was scoring a first or a second test. Each examiner first scored all tests he had given. The examiners then exchanged tests and each scored the tests of the other. Any items on which the two examiners failed to agree (there were only six of these in the 182 Stanford-Binets scored) were referred to another authority on the tests and the item scored in accordance with his decision. All tests were checked a final time by the author and rechecked by a competent assistant. All computations of M.A. and I.Q. were re-checked several times by the author and by a competent assistant.

The Binet Examiners

Dr. Mildred Stanton who assisted the author in this study has wide experience in the administration of mental tests. She has fifteen years of experience with Binet individual tests and much experience with the 1937 Revision forms L and M. She has much practice in administering individual intelligence tests to preschool age children.

The author of this research received training in the administration of the 1937 Revision of the Stanford-Binet Test at Teachers College in 1938. Since that time he has administered several hundred tests to public and private school children in New York City and its environs. Of these tests approximately one hundred and twenty-five were administered to preschool and kindergarten subjects before beginning this study.

Chapter IV

Results

Comparison of results from two examiners

Throughout the handling of the data the results from Dr. Stanton's group of 48 subjects and the writer's group of 43 subjects have been assembled separately and then combined. It will be remembered that Dr. Stanton tested her group of children without knowledge of the author's hypotheses or aims. No significant differences between the results from the two groups have been discovered. In terms of range of I.Q. change, Dr. Stanton's group showed changes ranging from a loss of 21 to a gain of 26 I.Q. points. The writer's group showed changes ranging from a loss of 13 to a gain of 28 I.Q. points. The two groups showed an approximately equal proportion of gains above 10 I.Q. points. Table 1 presents critical ratios for differences between the mean M.A. and I.Q. scores of the two groups together with C.R.'s for differences between mean C.A.'s at time of initial and final tests.

Table 1

Critical ratios for differences between mean scores of writer's subjects and those of Dr. Stanton

Means	C.R.
M.A's first test	.68
M.A's second test	.96
I.Q's first test	.80
I.Q's second test	.61
GAINS in I.Q. 1st to 2nd Test	2.02[1]
C.A's first test	.00
C.A's second test	.78

[1] Half of Dr. Stanton's group were retested after an interval of one month and the remaining half after two months. All of the author's group were retested after two and three month intervals.

From these data it is concluded that there are not sufficient differences in the results obtained by the two examiners to warrant separate treatment in this report.

The Stanford-Binet test scores

This section, Tables 2 to 7, contains the results of both initial and final testing of the 91 children on the Stanford Revision of the Binet Test, 1937 (37). The mean interval between tests is 1.93 months, and the mean school attendance between tests is 30.2 three-hour sessions, S.D. 12.24. (See Table 8.) The group consisted of 44 boys and 47 girls. 45 members of the group were initially tested on Form L of the test, and 46 were initially tested on Form M. Each child was retested on the form alternate to the form of his initial test. All measures of central tendency, all measures of dispersion, and all coefficients of correlation reported have been machine calculated from ungrouped data. Unless otherwise stated all correlation coefficients are Pearson coefficients. All results were checked by a second independent calculation to insure accuracy.

a) Chronological Age

A comparison between the C.A.'s of these subjects as reported in Table 2 and the C.A.'s of the subjects of Anderson (1), Bayley (4), Bird (6), Franden and Barlow (12), Goodenough and Maurer (18), Honzik (22), Lamson (25), Olson and Hughes (29), Starkweather and Roberts (25), Voas (39) and Wellman (9, 40 through 51), shows that this group is

TABLE 2
Initial and final C.A's

Months	Freq. 1st test	Freq. 2nd Test.
75–79	1	1
70–74	1	8
65–69	24	26
60–64	37	43
55–59	28	13
N	91	91
Range	56–75 mo.	57–76 mo.
Mdn	61 mo.	63 mo.
Mean	62.02 mo.	64 mo.
S.D.	4.38 mo.	3.97 mo.

somewhat more homogeneous as to C.A., and that its mean C.A. is slightly higher than the C.A.'s reported by these authors.

b) Mental Age

A correlation of $+.80\pm .03$ has been obtained between the mental age scores of tests one and two. The children made a mean gain of 5.84 months in mental age from test to re-test after a mean interval of 1.93 months during which time they had attended preschool a mean of 30 half day sessions. The critical ratio for the difference between the mean mental ages obtained is 9.90.[2]

TABLE 3
Initial and final Stanford-Binet mental age scores

Months	Freq. 1st Test	Freq. 2nd Test
90–99	0	1
80–89	5	5
70–79	15	31
60–69	29	37
50–59	37	15
40–49	5	2
N	91	91
Range	44–84 mo.	49–90 mo.
Mdn	61 mo.	67 mo.
Mean	62.09 mo.	67.93 mo.
S.D.	9.04 mo.	8.60 mo.

$r_{MA1-2} = +.80 \pm .03$
$CR_{MA1-2} = 9.90$

Table 3 below shows a distribution of the individual changes in M.A. from test to re-test.

TABLE 4
M.A. changes from test to re-test

M.A. Change Months	Freq.
+20+24	2
+15+19	4
+10+14	15
+ 5+ 9	32
0+ 4	30
− 5− 1	5
−10− 6	3
N	91
Range	−9 mos. to +21 mos.
Mdn	+6 mos.
Mean	+5.84 mos.
S.D.	5.53 mos.

c) I.Q.

TABLE 5
Initial and final Stanford-Binet I.Q.'s

I.Q's	Freq. 1st Test	Freq. 2nd Test
125–134	3	7
115–124	11	16
105–114	16	33
95–104	32	21
85– 94	21	11
75– 84	6	3
65– 74	2	0
N	91	91
Range	66–133	77–133
Mdn	100	107
Mean	100.11	106.18
S.D.	12.18	11.09

$r_{IQ1-2} = +.75 \pm .03$
$CR_{IQ1-2} = 6.98$

A correlation of $+.75\pm .03$ was obtained between the I.Q. scores of tests one and two. Wellman (42) obtained an

[2] C.R. taking r into account; see Garrett (13, p. 217)

r of .71± .05 between initial and re-test after an interval of 18 months administered to a group of 77 preschool children. Her initial test was administered at time of entrance to preschool and final test at the end of 18 months of preschool enrollment. Wellman states that she found the greatest consistency (51, p. 63) between the tests during this interval of greatest gain. Her subjects made a mean gain of 9.3 I.Q. points from initial to final test. However, the major gain of 8.6 points was made in the initial 18 months of attendance in the University system, and no real gain was made thereafter.

Direct comparisons may not be made between the correlation obtained in the present study and that obtained by Wellman, since the C.A. of her group ranged from 24 to 72 months with a mean of 48.2 months. It is also impossible to tell how many of her children were tested finally on the same test on which their initial I.Q's were established. She states that Kuhlmann-Binets were used with children up to 42 months of age and Stanford-Binets 1916 unrevised with children above 42 months of age. Wellman (51, p. 392), however, states that correlations she has obtained "between test and re-test I.Q. at preschool ages fall within the ranges of correlations obtained at school ages."

It might also be pointed out that the interval between tests in this study is much shorter than the interval between tests in other studies reported in the literature and that it is considerably longer than the "one to a few days interval" of Terman and Merrill.[3] Bayley (4, 5) and Honzik (22) have found higher correlations between tests administered closer together in time at the preschool ages and lower correlations for those separated by longer intervals. It will be necessary to have additional test results for this group before direct comparisons with the total data of these authors can be made. Honzik reports a correlation of .68 between test and re-test at the preschool level after a mean interval between tests of 1.90 months.

Bird (6), in a study in which 54 children (mean C.A. 56 months at time of initial test) were initially tested within one month after entering preschool and finally tested one month before the end of preschool experience on 1916 Stanford-Binet, obtained a correlation of +.84 between test and re-test I.Q.'s. The difference between Bird's result and the correlation of +.75±.03 reported here might conceivably be accounted for by pointing out that the tester who gathered Bird's data was reported to have been in close continuous contact with the children studied from the time of their entrance at preschool until the final tests were administered. Bird's children were not initially tested until after some preschool experience had been obtained and, in view of the above statement, probably not until a majority of them were acquainted with the examiner in their preschool routine.

The possibility of obtaining a more reliable I.Q. when the initial I.Q. is obtained after some school experience has been pointed to by Updegraff (38), who obtained an r of +.58± .06 between the intelligence scores of a group of 63 subjects who were initially tested during the week previous to the opening of the Iowa Preschool Laboratory and finally tested at the end of their first year of preschool experience. However, with another group of 123 subjects

[3] Terman and Merrill (37, p. 47) report that the median reliability for Stanford-Binet tests 1937 Revision, correlating Form L against Form M for ages 2 to 6 years is .88.

of the same age range, initially tested after at least two weeks and not more than two months of preschool experience, Updegraff obtained a correlation between first and second tests of $+.84 \pm .02$. Updegraff's subjects were initially and finally tested on the same intelligence test, the younger members of the group being tested on the Kuhlmann-Binet and the older members on the 1916 Stanford-Binet. In discussing her study Updegraff (38, p. 158) states:

> To assume that the results of studies of constancy of the I.Q. in older children would be indicative of the outcome when similar methods are applied to the young child would be unjustifiable; the differences in the standardization of the tests at the preschool ages indicates inaccuracy of the testing medium, not to speak of the emotional factors which may have, if not a greater, at least a different influence on the young child in the test situation.

The critical ratio for reliability between the mean I.Q.'s obtained on the two tests in the present investigation is 6.98.

Table 6 below shows a distribution of the individual changes in I.Q. from test to re-test.

TABLE 6
I.Q. changes from test to re-test

I.Q. Changes	Freq.
+25+29	2
+20+24	4
+15+19	7
+10+14	20
+ 5+ 9	17
0+ 4	23
− 5− 1	12
−10− 6	2
−15−11	3
−20−16	0
−25−21	1

N = 91
Range = −21 to +28
Mdn = +6
Mean = +6.07
S.D. = 8.67

Tables 5 and 6 show that 91 children initially testing at a mean of 100.11 I.Q. before preschool experience gained 6.07 I.Q. points when re-tested after a mean interval of 1.93 months. During this interval they attended preschool a mean of 30.2 half day sessions (Table 8). This mean gain of 6.07 I.Q. points from test to re-test compares favorably with that of Wellman (51, p. 382) who states: "The 652 individuals who attended preschool for one year made a gain of 6.6 points between the fall and spring tests. Their mean initial I.Q. was 116.9; changes were widely distributed." In a table (51, p. 382) it is shown that changes in I.Q. for this group of 652 subjects range from losses as great as 37 to gains as great as 47 I.Q. points. The group of 91 subjects reported on in Table 6 of the present study has a range of changes in I.Q. from a loss of 21 to a gain of 28 points. With less than one-seventh of the subjects the range of I.Q. changes is more than one-half that of the Wellman group. The results reported here deal with subjects more homogeneous as to C.A. (Table 1); while the Wellman data (51, p. 381) deal with subjects ranging from 18 to 77 months in C.A. Wellman states (46):

> We find year after year increases in the I.Q. of children enrolled in Iowa University nursery schools. In general for the whole preschool population in any given year the increase in I.Q. is approximately 5-8 points. A greater gain is made by children who attend preschool for two consecutive years.

In (44) she states further:

> relationship between the child's I.Q. and environmental forces is affected by the amount of change that the environment brings forth. Distance (amount of change) and time are necessarily interrelated. In general, greater changes may be expected over longer periods of time.

and again in (50) "In order that change may take place, it is necessary to allow some time for the change to occur," and from (41) "Significantly greater gains were made by those children having seven or eight months of preschool attendance between their tests than by those having five or six month intervals." From these statements plus the report of a mean change of 6.6 I.Q. points after one year of attendance at the University of Iowa Nursery Schools by 652 children, the writer believes himself to be correct in interpreting Wellman to mean that this change in I.Q. occurred as a result of at least a year's attendance at preschool. If this interpretation is correct, would it not be just as correct to say that the 91 subjects reported in this study profited equally well, in so far as mean increase in intellectual capacity is concerned, from a mean of 30 half day sessions of a public school kindergarten? It is not the intention to interpret the gain in I.Q. here demonstrated as an increase in intellectual capacity. The change has occurred so rapidly as to preclude such an interpretation. A presentation and analysis of additional data plus further analysis of data already presented will suggest an explanation of I.Q. changes reported here in terms of adjustment of the child to the testing situation.

In Wellman's 1932 article (40) we find it reported that 1,333 children had gained a mean of 9 points in I.Q. from initial testing to an average of their second and third tests with an interval of approximately twelve months between tests. From a chart on page 462 of this study it is possible to approximate the amounts by which each age level contributed to this mean gain of 9 points. Children three years old at the time of the initial test, of whom there were 172, appear from the chart to have gained an average of 23 I.Q. points from test to re-test. Children who were four years old at the time of the initial test, N = 150, appear to have gained an average of 25 I.Q. points from test to re-test, while children 5 years old at the time of the initial test, N = 176, appear to have gained only 5 to 6 I.Q. points from test to re-test. From this report it seems reasonable to argue that Wellman's gain of 8.6 I.Q. points reported in a later article (42) for preschool attendance only may have been considerably contributed to by the younger members of her group. It also might be argued that it appears as if Wellman's greatest gains occur in the age area in which the subjects are more likely to have been tested initially on the Kuhlmann-Binet and finally on the Stanford-Binet. It is also interesting to note that the mean gain of 6.07 I.Q. points by children with mean C.A. of 62 months reported in the present study slightly exceeds the 5 to 6 points gain indicated by Wellman (40) for 176 children initially tested at a C.A. of five years. Both Honzik (22, 23) and Anderson (2) have pointed out that mental test constancy appears to be markedly dependent upon the ages of the children at the time of testing.

Several other research workers have also reported gains in I.Q.'s of preschool age children after a period of preschool experience. Bird (6), studying 54 subjects with a median C.A. of 56 months, reports a mean gain in I.Q. of 1.8 points after approximately one year of preschool experience. Since Bird's examiner was in close continuous contact with the children studied and since each child was initially tested within one month after entrance to preschool, it is very probable that the children were better acquainted

with the examiner than is usual in most testing situations. It is possible that the mean gain of 6.07 I.Q. points reported in the present study, may be partially accounted for as better adjustment by the children, after some preschool experience, to the teacher-pupil situation involved in the administration of the Stanford-Binet Test, to greater familiarity with the surroundings in which the test was administered and to what might be called a certain degree of test wisdom, having once before come through one such experience without damage. This tentative explanation appears to have a considerable amount of support in the researches of Caille (7), Reynolds (30), Rust (31), Levy and Tulchin (26), Jersild and Fite (24), Updegraff (38) and Hallowell (20); and in the opinions of Anderson (3), Bayley (5), Cattell (8), Dearborn and Rothney (10), Goodenough (18), Hollingworth (21), and many other psychologists and educators.

Frandsen and Barlow (12), using Form L of the 1937 Revision of the Stanford-Binet for both initial and re-test after approximately 5½ months of preschool experience for 30 children with a mean C.A. of 43.61 months, report a mean gain of 3.34 I.Q. points from test to re-test with a C.R. of 2.81. These authors conclude that the gain, although it approximates statistical significance, appears very small when compared with individual differences resulting from both hereditary and environmental causes. In this research, as in that of Bird (6) above, the initial tests were administered during "approximately" the first month of preschool experience.

Anderson (3) reports a mean gain of 2.63 I.Q. points for 26 children initially tested "in most instances" after some preschool experience and re-tested after a mean interval of approximately six months of preschool experience.

A consideration of the data in Table 7 below, with respect to change in I.Q. according to initial I.Q., in comparison with reports of this nature published by Wellman and others, will not be attempted here. The data indicate that subjects initially testing low made greater gains in I.Q. on the second test than those testing near or above average and that those testing high suffered losses in score. This relationship between gains and initial scores is to be interpreted as a necessary effect of the unreliability of the initial measurement, as a tendency for individual scores to regress towards their true means rather than "a tendency for changes in I.Q. to be related to the general I.Q. level of the group" (27, p. 84).

TABLE 7

Distribution of I.Q. changes according to level of initial I.Q.

Initial I.Q. Level	No. of Children	Mean I.Q. Points Change 1st to 2nd test
125–134	3	−5.00
115–124	11	−1.18
105–114	16	+3.00
95–104	32	+8.03
85– 94	16	+9.05
75– 84	6	+9.17
65– 74	2	+15.00

Section 2 has shown that the 91 children studied did make significant gains in I.Q. after a period of preschool experience. The mean gain of 6.07 points was obtained after a mean attendance of 30 three-hour sessions at preschool and after a mean interval of 1.93 months between the tests. The short period of preschool experience and the short interval between tests preclude the interpretation of gains obtained as resulting from mental growth due to preschool experience.

DATA RELATED TO THE BINET TEST SCORES

a) Length of Pre-school Attendance

Table 8 below shows a distribution of the numbers of three-hour sessions of public school kindergarten attended by the subjects during the interval between tests.

TABLE 8
Three-hour kindergarten sessions attended

Sessions attended	Freq.
45–49	6
40–44	21
35–39	14
30–34	17
25–29	7
20–24	0
15–19	5
10–14	20
5– 9	1
N	91
Range	6 to 49
Mdn	34
Mean	30.20
S.D.	12.24

A comparison of the extreme groups in Table 8—26 children who attended kindergarten from 5 to 19 three-hour sessions with 27 children who attended from 40 to 49 sessions—shows that the short attendance group gained a mean of 3.73 while the long attendance group gained a mean of 6.63 I.Q. points. The correlation between three-hour sessions of kindergarten attended and changes in I.Q. was found to be $+.15 \pm .06$. These data appear to support those reported in Tables 9 and 10 on changes in I.Q. in relation to interval between tests.

b) Interval between Tests

The mean interval between tests was 1.93 months. Twenty-six of the children were re-tested after an interval of exactly one month, 44 after two months, and 21 after three months. Table 9 below shows the means and S.D's of the initial and final I.Q's for these three groups. Table 10 shows C.R.'s for differences between these means.

If these I.Q. gains could all be taken at face value, we could conclude that the first month in preschool gave a gain of 4.31 points, the second month a gain of $6.86 - 4.31 = 2.55$ points, and the third month no further gain. But the C.R.'s in Table 10 show that the differences between groups are unreliable, though the superiority in gain of the two month's interval group over the one month shows a slight approach to significance. This trend, plus the data from Table 8, seems to indicate that the gain in I.Q. is to some small degree related to the

TABLE 9
Initial and final Binet I.Q's for groups retested after 1, 2, and 3 months.

Group	Meas.	I.Q. 1	I.Q. 2	Diffs. I.Q.
1 month Interval	Mean	99.46	103.77	+4.31
N=26	S.D. dis.	13.00	12.04	6.97
2 month Interval	Mean	99.20	106.06	+6.86
N=44	S.D. dis.	11.82	10.77	7.77
3 month Interval	Mean	102.81	109.38	+6.57
N=21	S.D. dis.	13.64	11.54	11.59
2 to 3 months Interval	Mean	100.37	107.14	+6.77
N=65	S.D. dis.	12.55	11.13	9.18

length of the preschool experience. If we should accept Wellman's findings of a mean gain of 6.6 I.Q. points (51) after one year of preschool experience as a criterion of the worth, in terms of intellectual growth, of one year of such experience to the average child, we could conclude that the children here studied are extremely fortunate in their selection of a school since their mean intellectual growth approximates the criterion after a mean of 30 three-hour sessions of school attendance. It seems more logical to believe that the gains reported in the present study are a result of adjustment, or are due, at least partially, to lack of adjustment at the time of the initial tests to the situation which the individual intelligence test involves. The trend in the data presented in Table 9 plus the data from Table 8 may indicate that the time necessary to make an adjustment sufficient to register mean positive changes in an upward direction in I.Q. comparable to those attributed by Wellman to a year of preschool experience probably is less than two months of such experience. The small coefficient of correlation .15 ± .06 for length of attendance versus I.Q. change might be used to support the argument that the length of time necessary up to two months would depend to a great extent upon the capacity for adjustment of the individual child.

TABLE 10
C.R.'s for differences between mean I.Q.'s of groups retested after 1, 2, and 3 months

Groups	I.Q.'s 1st Test	I.Q.'s 2nd Test	Diffs. I.Q.'s 1–2 Test
Means of 1 mo. Group vs. means of 2 mo. group	.08	.80	1.42
Means of 1 mo. Group vs. means of 3 mo. group	.85	1.63	.78
Means of 2 mo. Group vs. means of 3 mo. group	1.04	1.11	.10
Means of 1 mo. Group vs. means of 2–3 mo. group (Combined)	.30	1.23	1.39

Several other points have been brought out by a study of the Binet Test results. It was found that 44 boys earned a mean change of +6.45 I.Q. points from test to re-test, while 47 girls earned a mean change of +5.70 I.Q. points from test to re-test. Of the 44 boys studied 20 were initially tested on Form L and 24 on Form M. Of the girls studied 25 were initially tested on Form L and 22 on Form M. No plans were made in advance as to what form should be used with any particular child. Each examiner was equipped with both Form L and Form M materials and tested the children as they arrived, alternating between forms so as to test half of the group initially on Form L and half on Form M. Another point of interest is that 45 children (25 boys and 20 girls) initially tested on Form L made a mean gain on re-test on Form M of 4.62 I.Q. points, while 46 children initially tested on Form M (24 boys and 22 girls) made a mean gain on re-test on Form L of 7.67 I.Q. points. According to Terman and Merrill (37, p. 43) "The mean increase in I.Q. on the second test is the same from L to M as from M to L. The amount of increase shows no noticeable trend in relation to size of I.Q. but varies according to the age of the subjects." These authors report that the mean practice effect for ages 5 to 16 years is approximately 2 I.Q. points. This effect was obtained when the tests were administered within an interval of from one to a few days and

the authors caution against inferring the same amount of practice effect when the tests are separated by a longer time interval. These authors further state (p. 43),

> Extensive data from re-tests with the old Stanford-Binet indicate that when the same scale is repeated, a small effect may persist for several months, and that in the case of subjects who have been given the test several times, the practice effect may become a more or less serious matter. If the two scales now available are used alternately, a subject may be re-tested at relatively brief intervals without risking any vitiation of the results by reason of practice effects.

On page 40, Terman and Merrill give the variability of the two forms of the test in relation to C.A. of subjects tested. They have found the S.D. of L for five-year-old subjects to be 14.2 and for M with the same subjects to be 14.1. In view of this report and that above, the differences in gains from L to M and from M to L ought not to have occurred. It is suggested that this difference may be explained as resulting from lack of as much practice with Form M on the part of the examiners as had been obtained by them on Form L. This explanation does not seem to be too likely since the gains run as they do. It is possible that differences in familiarity with the two tests may have operated to depress initial M scores allowing for a slightly greater gain in I.Q. from M to L. However, the two examiners, Stanton and McHugh, were well experienced in both forms of the test prior to the beginning of the study; and lack of familiarity with Form M on the second test, after the two examiners had administered 45 of these at this age range on the initial test, should not have operated to depress the final M I.Q.'s making for a smaller gain for the group who were tested first on Form L and second on Form M.

c) C.A. and I.Q. Changes

Table 11 below shows the relationship between C.A. at the time of the initial test and gain in I.Q. points first to second test.

A correlation of $-.053 \pm .067$ was obtained between change in I.Q. 1st to 2nd test, and C.A. at the time of the first test. These data indicate no significant relationship for this group over this age range between C.A. at the time of the initial test and I.Q. gain 1st to 2nd test. In another research project, these children are being re-tested after a year of kindergarten experience. The new data will be studied further with reference to the relationship reported here.

TABLE 11
Gains in I.Q. second test in relation to C.A. at time of first test

1st Test C.A. Mo.	No.	Mean I.Q. gain 1–2 test
75–79	1	4.00
70–74	1	26.00
65–69	24	4.17
60–64	37	5.18
55–59	28	7.39

Data presented in section 3 do not warrant concluding that there is a significant relationship between length of preschool experience and gain in I.Q., or that any significant relationship exists between C.A. at the time of the first test and gain in I.Q. after preschool experience for the 91 children studied under the conditions of this research.

Evidence offered in this section supports the hypothesis that gains in I.Q. at the preschool level are adjustmental rather than a result of growth in mental capacity. The subjects' gain after a mean attendance of 30 sessions at preschool is comparable to the mean gain of Wellman's subjects (51, p. 382) after a year's

attendance at preschool. Wellman attributes her subjects' gain to growth in intelligence due to a year of preschool experience.

FURTHER ANALYSIS OF THE BINET RESULTS

a) Basal and Ceiling Ages

An attempt is made here to compare the basal ages earned on the first and second Stanford-Binet Tests, and to compare the ceiling ages earned on the first and second tests. Comparisons are also made between changes in basal age from first to second tests and changes in ceiling ages first to second tests. Basal and ceiling ages are defined by Terman and Merrill (37, p. 63) as that age level at which all test items are passed, and that age level at which all test items are failed, respectively.

Tables 12 and 13 below show distributions of basal age scores first and second tests, and ceiling age scores first and second tests.

TABLE 12
Basal ages in M.A. months

Basal Age M.A. Months	Frequency 1st test	Frequency 2nd test
72	3	8
60	17	26
54	18	22
48	32	28
42	14	6
36	7	1
N	91	91
Range	36–72 mo.	36–72 mo.
Mdn	48	54
Mean	50.37	54.46
S.D.	7.99	7.92
CR = 2.54		

A correlation of $+.48 \pm .05$ was obtained between basal ages first test and basal ages second test.

A C.R. of 2.54 was obtained for reliability of difference between mean basal age first test and mean basal age second test. This C.R. was obtained by first subtracting 1.93 months from the difference obtained, on the assumption that since a mean of 1.93 months had elapsed between these tests one would expect at least 1.93 months change in basal age from test to re-test.

A correlation of $+.45 \pm .06$ was obtained between ceiling ages first test and ceiling ages second test.

A C.R. of 4.24 was obtained for reliability of difference between mean ceiling age first test and mean ceiling age second test. As in the case of the difference between basal age means, 1.93 was subtracted from the difference obtained before calculating the C.R.

TABLE 13
Ceiling ages in M.A. months

Ceiling age M.A. Months	Frequency 1st test	Frequency 2nd test
132	0	2
120	0	1
108	15	24
96	12	26
84	42	30
72	20	8
54	2	0
N	91	91
Range	54–108 mo.	72–132 mo.
Mdn	84	96
Mean	86.24	94.15
S.D.	12.66	12.94
CR = 4.24		

Tables 12 and 13 show that the children studied made mean and median gains in ceiling age approximately twice as great as the mean and median gains in basal age.

In an effort to determine relationships between these changes in basal and ceiling ages and obtained changes in I.Q. from test to re-test, correlations were computed between change in basal age from test to re-test versus change in I.Q.,

and change in ceiling age from test to re-test versus change in I.Q. A correlation coefficient of .33 ± .06 was obtained between basal age changes and I.Q. changes and a correlation coefficient of .64 ± .04 between ceiling age changes and I.Q. changes. The C.R. of the difference between these coefficients, D/PE_d, is 4.43, indicating a significant and reliable difference. One may conclude from these data that the increase of I.Q. from test to re-test of the children studied is more closely related to increase in ceiling age than to increase in basal age. A further analysis of the individual test items presented below may make it possible to infer that this relationship exists because of 1) the nature of the test over the age range covered by these children, and 2) the nature of the changes in the child due to a short period of pre-school experience.

b) Individual Test Items

The lowest age level of the Stanford-Binet 1937 Revision Form L or M administered in this study was 3 years. The highest age level administered was 11 years. This analysis, then, concerns itself with individual test items on both forms of the test that fall between the 3 and 11 year age levels. A survey of the test items with the view of classifying them by type of item reveals that all items of Form L and all but two of Form M in the age levels of 3 through 11 years may be classified into two distinct categories: 1) Items that can be successfully passed by the child through some manual or manipulative behavior and without the use of oral speech, 2) Items that cannot be successfully passed by the child without the use of oral speech. Table 14 below shows this classification in which "S" indicates an item requiring oral speech, "M" an item not requiring oral speech and "D" an item that cannot be classified.

TABLE 14
Classification of Binet items

Yr. Level	Form L Items							Form M Items						
	1	2	3	4	5	6	Alt.	1	2	3	4	5	6	Alt.
III	M	S	M	M	M	S	M	M	S	M	M	S	S	M
III-6	M	S	M	S	M	S	M	M	M	M	S	M	S	M
IV	S	S	M	M	M	S	S	S	M	S	M	S	S	M
IV-6	M	S	M	S	M	S	M	M	S	S	M	S	S	M
V	M	M	S	M	S	S	M	S	M	D	M	S	S	M
VI	S	M	S	M	M	M		M	M	S	S	D	S	
VII	S	S	M	S	S	S		S	S	S	S	S	S	
VIII	S	S	S	S	S	S		S	S	S	S	S	S	
IX	M	S	M	S	S	S		M	S	S	S	S	S	
X	S	S	S	S	S	S		S	S	S	S	S	S	
XI	M	S	S	S	S	S		S	M	S	S	S	S	

Although alternate test items are presented in Table 14 they were not used in the further analysis of the individual results. Since alternate items were used only four times throughout the testing they are omitted from the data below.

Table 14 shows that the 36 items, excluding alternates, in Form L from the 3 year level through the 6 year level are made up of 20 manual items and 16 speech items, and that these two types of items occur about 50% of the time at each age level. From the seven-year level through the 11-year level, however, the

30 items of Form L are made up of only 4 manual items as against 26 speech items.

The 36 items of Form M from 3- through 6-year levels are made up of 15 manual items, 19 speech items and 2 doubtful items, and the 30 items of Form M from 7- through 11-year levels are made up of 2 manual items and 28 speech items.

Summarizing in terms of per cents of manual and speech items we have:

Year Levels	Form L		Form M		
	% Manual	% Speech	% Manual	% Speech	% Doubtful
3–6 incl.	44	56	42	53	5
7–11 incl.	13	87	7	93	

It is apparent that both forms L and M require a great deal more oral speech behavior of the testee above the six year level than they do up through that level. Because of the sudden change in the type of items in the direction of requiring more speech behavior above the six year level comparison has been made here of individual changes from test to re-test over the range of the 3 through 6 year levels only. Within this range the children had about equal opportunities to attempt manual and speech items. An analysis has been made of individual test items according to their classification (Table 14) and according as the individual item was passed, failed or refused by the individual testee. An item was considered to be refused when the testee failed to react to it in any way or when the testee answered, "I don't know," and refused to attempt to pass the item after urging (where the instructions of the test allowed urging). Table 15 below presents a summary of findings obtained by the method of analysis described above.

Table 15 shows that with approximately equal opportunity to improve in both manual and speech items from first to second test and over the same age levels of the test this group of 91 children showed 11.2% improvement in speech items as compared with 4.7% in manual items. It will also be noted that the group refused many more speech items on the first test than manual items. From these data one may infer that the mean gain of 6.07 I.Q. points obtained

TABLE 15

Results of analyses of individual Binet items 3 through 6 year levels first and second tests

Speech Items 3–6 yr. level, Forms L and M:			Total Speech Items	% Speech Items	% Speech Passed	% Speech Failed	% Speech Refused
Children	Test	Total Items Presented					
91	1	1440	712	49.4	54.3	38.8	6.9
91	2	1098	530	48.3	65.5	31.5	3.0
Manual Items 3–6 yr. level, Forms L and M:			**Total Manual Items**	**% Manual Items**	**% Manual Passed**	**% Manual Failed**	**% Manual Refused**
Children	Test	Total Items Presented					
91	1	1440	650	45.2	51.4	46.9	1.7
91	2	1098	499	45.4	56.1	43.7	.2

from 1st to 2nd tests is more closely related to an improvement in oral speech behavior during the interval between tests than to a general increase in intellectual ability. It is impossible to say whether this improvement is due to learning to use the tool of language in school or to a reduction of factors that may have operated to inhibit oral speech behavior at the time of the first test. Such factors might be: inexperience in talking to teachers, shyness in a new situation, or actual fear of the examiner who was a strange person. In view of the short period of school experience the proposed explanation in terms of reduction of factors that may have operated to inhibit oral speech behavior during the first test seems more acceptable.

The situation becomes clearer if we present the same analysis of the test results of 23 children in this study who gained a mean of 16.83 (range 12 to 28) from test to re-test, as compared with a similar analysis of the test results of 38 children who gained a mean of .79 (range of +5 to −5) I.Q. points.

From Table 16, 23 children who made large gains in I.Q. from test to re-test improved 24.4% in speech items passed on the second test, as compared with an improvement of 1.9% in the group of 38 children whose second I.Q.'s were within ±5 points of the initial I.Q.'s. The group of 23 children gaining in I.Q. also improved 12.1% in manual items passed from test to re-test as compared with a loss of 2% in manual items passed by the non-gaining group of 38 children. It will also be noted that the gaining group refused 10.2% of speech items presented at the first test as compared with 3% refused by the non-gaining group. These data further support the inference that gains in I.Q. from test to re-test of preschool age children may be, at least partially, accounted for in terms of improvement in the use of oral speech.

Tables 15 and 16 also support a possible explanation of the significantly

TABLE 16
Results of analysis of individual Binet items 3 through 6-yr. levels for 23 children making large I.Q. gains and 38 children making no gains.

Speech Items 3–6 year level inclusive. Forms L and M.

Children	Test	Total Items Presented	Total Speech Items	% Speech Items	% Speech Passed	% Speech Failed	% Speech Refused
23G	1	407	213	52.3	46.9	42.3	10.2
23G	2	252	115	45.8	71.3	23.5	5.2
38NG	1	591	288	48.7	59.7	34.7	5.6
38NG	2	487	242	49.7	61.6	36.8	1.6

Manual Items 3–6 year level inclusive. Forms L and M.

Children	Test	Total Items Presented	Total Manual Items	% Manual Items	% Manual Passed	% Manual Failed	% Manual Refused
23G	1	407	166	40.8	44.6	52.4	3.0
23G	2	252	127	50.6	56.7	43.3	0.0
38NG	1	591	276	46.7	54.3	44.6	1.1
38NG	2	487	212	43.5	52.3	47.2	.5

closer relationship of gains in ceiling age to I.Q. gains, than the gains in basal age to I.Q. gains. Since the Stanford-Binet test Form L is made up of 87% and Form M of 93% of items requiring oral speech for the 7 to 11 year age levels, and since 89 of the 91 children earned initial ceiling ages of 6 years or better (Table 13) almost all improvement in ceiling would necessarily be at the age levels requiring oral speech behavior to an extreme degree. It seems reasonable to believe that improvement in ceiling age for this group is related to improvement in I.Q. through improvement in oral speech behavior, and that this relationship explains the significant difference obtained between correlations for ceiling gains versus I.Q. gains and basal gains versus I.Q. gains.

Section 4 supports the hypothesis that gains in I.Q. after preschool experience are primarily adjustmental and are not due to growth in intellectual capacity. It is suggested that one aspect of the adjustment is improvement in the use of oral speech in the testing situation.

THE EXAMINER'S JUDGMENTS OF RAPPORT

a) Spontaneous Verbalization

The examiners recorded at the beginning of each Binet test a judgment of the amount of spontaneous verbalization made by each child during the period prior to beginning the test. These judgments were recorded on the child's record blank as:

1. None
2. Little
3. Moderate
4. Free
5. Incessant

It was felt that verbalization might serve as a kind of behavioral index of the amount of shyness felt by the child at the beginning of the test and the hypothesis was suggested that an increase of spontaneous talk at the time of the second test would be related to gains in I.Q. from test to re-test. An analysis of the results obtained from this approach reveals that 53 of the 91 children made no spontaneous remarks prior to the beginning of the initial test, 20 subjects were placed in the category of "little" at the first test, 12 in the category of "moderate," and 6 in the category of "free." On the second test the numbers in each category were as follows: None, 11; Little, 17; Moderate, 32; Free, 29; Incessant, 2. The entire group moved a mean of 1.23 categories in the direction of more "talkativeness" prior to the beginning of the second test. An analysis of the individual results, however, reveals no relation between increased spontaneous verbalization at the beginning of the second test and gain in I.Q. score on the second test. The 18 subjects who lost in I.Q. from 1 to 21 points on the second test changed an average of 1.61 categories in the direction of more talking at the time of the second test; the 11 children who registered zero change in I.Q. made a positive change of 1.57 categories, while the 66 children who gained in I.Q. from 1 to 28 points (I.Q.) made a positive change of only 1.09 categories. And the 33 children who gained 10 I.Q. points or more also made a mean positive change of only 1.09 categories. A further analysis shows that the 68 children making positive changes in amount of spontaneous verbalization gained an average of 5.71 I.Q. points, that 7 children making negative changes gained 8.71 I.Q. points; and that 16 children making no changes gained 6.44 I.Q. points.

Contrary to our hypothesis, then, I.Q.

gain was not the result of increased freedom of speech at the beginning of the test session.

b) Shyness and Negativism

Scales for shyness and negativism already described on page 16 were used. The examiner marked off on these scales, at the end of each testing period a line which indicated the extent to which, according to his judgment, the child had manifested these traits during the testing period. Precautions were taken to file away the records of the first test so that all second test judgments were made without reference to the first test judgments. Individual scores for these traits were obtained by measuring off in millimeters from the right end of each scale to the point marked off by the examiner, high scores thus indicating the manifestation of extreme shyness and negativism and low scores indicating an absence of these traits, as judged by the examiners. No attempt was made to standardize the scales before using them. It was hoped that cases manifesting extreme shyness or negativism would have so clearly changed by the time of the second test that the examiners would have no difficulty in observing and recording the change.

Since two examiners made judgments on these scales on different groups of children there is no way to check one examiner's judgments against the other; at no time were the two examiners' scores on these traits treated together until these scores had been converted to standard scores with reference to their particular group means (13, p. 178).

Table 17 below was constructed after an analysis of shyness and negativism scores assigned to the children during initial and final tests had been made.

It might be expected that the largest gains in I.Q. would be made by children originally the shyest—and also by the children who were finally least shy, but from Table 17 it appears not only that those children who were judged to be most shy and most negativistic at the time of the *first* test gained more in mean I.Q. points from first to second test, but also, that those children judged to be most shy and negativistic at the time of the *second* test gained more in mean I.Q. points from first to second test. It is possible that this finding was to have been expected in view of the way in which the examiners' judgments of shyness and negativism were made and recorded. It has already been stated that judgments on these traits at the time of the second test were made without reference to previous judgments. It might be pointed

TABLE 17

Mean gains in I.Q. from test to re-test as related to shyness and negativism

Test No.	Above Mdn Shyness	Below Mdn Shyness	Q_4 Most Shy	Q_1 Least Shy	Ten Most Shy	Ten Least Shy
1	6.96	5.31	6.91	4.91	10.60	5.80
2	8.47	3.76	8.57	1.87	6.70	1.60
	Negativism and mean gains in I.Q. points test to retest					
Test	Above Mdn Negativ.	Below Mdn Negativ.	Q_4 Most Neg.	Q_1 Least Neg.	Ten Most Neg.	Ten Least Neg.
1	7.80	4.38	9.09	2.70	9.30	3.0
2	7.27	4.91	6.78	4.78	7.50	1.20

out, further, that the judgments at the time of either test probably were influenced by the behavior of the group as a whole at the time of that test. No behavioral indices of shyness or negativism were recorded so that the examiners' judgment is in no way anchored to points along the scale. It might also be stated that one would not expect extreme changes for individuals of the group. A child judged to be shy or negativistic in relation to the group tested at the time of the first test would be likely to continue to be judged shy or negativistic in relation to the same group re-tested after an interval of time. In spite of this stability of position in the group the individual could have made some progress in the direction of becoming less shy or less negativistic during the period between tests. The principal difficulty with this type of approach lies in the difficulty of evaluating small positive or negative changes for the extremely shy and negativistic children, in terms of similar changes for children who manifest these traits to a much less degree from the very beginning of the study. As such scales are now constructed changes of equal magnitude are of equal arithmetical value even when made on contrasting ends of a scale.

Table 18 above shows the relationship between changes in shyness and negativism from first to second test as judged by the examiners and mean changes in I.Q. from first to second test.

In addition to the data presented above in Tables 17 and 18, coefficients of correlation have been obtained between change in the direction of becoming less shy and change in I.Q. from test to re-test, and change in the direction of becoming less negativistic and change in I.Q. In the change in shyness and I.Q. change a correlation of $+.05 \pm .07$ was obtained while in the case of change in negativism and I.Q. change a correlation of $.14 \pm .07$ was obtained.

TABLE 18

Mean I.Q. gain in relation to change in shyness and negativism scores first to second test[4]

	Shyness Change					
	Above Mdn	Below Mdn	Q_4	Q_1	10 largest	10 smallest
Mean I.Q. gain	7.18	5.04	8.13	5.61	6.6	6.6
	Negativism Change					
Mean I.Q. Gain	7.09	5.09	8.22	3.22	8.9	5.3

[4] Change is difference, in the direction of being less shy or negativistic, between examiners' ratings.

The data from Tables 17, 18 and the correlation coefficients above offer no acceptable evidence of a relationship between a decrease in shyness and negativism and increases in I.Q. for this group over the period of time studied. A most liberal interpretation of these results would allow for the possibility of a trend in the direction of a relationship between decrease in shyness and negativism and increase in I.Q. It is felt that the results obtained here demonstrate the inadvisability of concluding, from the individual examiners' estimation of the preschool child's behavior during the test period, that a large change in the I.Q. obtained can be expected on re-test

under better conditions of rapport with regard to shyness and negativism. While no proof of a relationship between shyness and negativism changes and I.Q. changes from test to re-test was found here, it is believed in view of the findings of Caille (7), Jersild and Fite (24), Rust (31) Updegraff (38), and others, that such a relationship may exist. It is suggested that future research might attack this problem by having trained observers rate the children in terms of rating scales on which points have been carefully anchored in terms of behavioral indices of these two traits.

TABLE 19
Vineland social quotients

Social Quotients	F
170–179	1
160–169	2
150–159	5
140–149	14
130–139	21
120–129	20
110–119	12
100–109	11
90– 99	2
N	88
Range	90 to 177
Mdn	129.30
Mean	129.30
S.D.	16.45

Section 5 offers no conclusive evidence to support the hypothesis that gains in I.Q. after preschool experience are related to improvement in the examiner's rapport with the child.

THE SOCIAL QUOTIENT

The Vineland Social Maturity Scale (11) was administered through a home interview with the mothers of the children in 88 cases. The scale was not completed in 3 cases because of inability to see the mothers after several visits.

Table 19 shows that a large majority of the children were credited by their mothers with social accomplishments considerably above the means for their C.A. group according to Doll's standardization of the scale (11). Since this scale has been used very little with normal unselected populations of preschool age level it is possible that the high scores obtained here are partially due to faulty standardization of the test at this age level. The writer recognizes, however, that the administration of this test requires a considerable amount of experience with it and that these high scores may result from faulty interview technique. It is suggested that in so far as a test of the Vineland Social Maturity type reflects environmental opportunity of the testee and in so far as such environmental opportunity operates to depress or increase the initial I.Q. of the preschool child, the S.Q. should correlate more highly with the initial I.Q. (obtained before preschool experience) than with the I.Q. obtained after some preschool experience. In the case of the 88 children reported here a correlation of $+.26 \pm .06$ was obtained between initial I.Q.'s and Vineland S.Q.'s, a correlation of $+.28 \pm .06$ was obtained between the second test I.Q's and these same Vineland S.Q.'s, and a correlation of $-.01 \pm .07$ was obtained between Vineland S.Q.'s and Binet I.Q. changes from test to re-test. These correlations offer no support for the suggestion above.

Contrary to our hypothesis, then, I.Q. gain was not related to degree of social maturity as measured by the Vineland test.

SOCIO-ECONOMIC RATINGS OF PARENTS

In addition to the tests administered to the children use was made of two standard rating scales to obtain socio-economic scores or ratings for each child. The two

scales were the Barr Occupation Rating Scale (36, p. 66) for the fathers' occupations and the Whittier Home Rating Scale (52 and 46, p. 73). Data for the Barr Scale consist only of the name given to the father's occupation by the mother at the time of the home interview. This information lists 50 different occupations for the 83 fathers reported. In five instances out of 91 the information was refused and in 3 it was impossible to contact either parent for a home visit. Those occupations named which directly fitted the classifications of the Barr Scale (36, p. 66) were rated directly on the scale; the remaining occupations were listed and submitted to 10 psychologists on the staff of Columbia University with a request that these individuals separately rate each occupation in terms of the Barr Scale. The final rating given each occupation was the mean of these ten separate ratings. The Whittier Home Rating Scale was used in accordance with the manual of instruction which accompanies the scale. Home ratings were obtained for 86 of the 91 children. These data were gathered by one examiner. In addition to the ratings above, educational information, the number of years of school attended, was obtained for both parents in 85 out of 91 cases. Cf. Table 20.

Table 21 shows a trend in the direction of positive correlation between socio-economic status of parents and the mental measures of children, but no significant relationship between socio-economic status and I.Q. change from first to second test. These findings are in agreement with those of Wellman (51), Bayley (5), Bird (6), Coffey and Wellman (9), Goodenough and Maurer (18), and others.

The writer would argue that results showing no significant relationship between environmental background at the beginning of preschool experience and

TABLE 20

Significant statistical data from measures of socio-economic status of parents at time of children's entrance to pre-school.

	Barr Scale	Whittier Scale	Years of Education Father	Years of Education Mother	Years of Education Midparent
N	83	86	85	85	85
Range	5.87–15.42	10–25	7–16	9–16	8–16
Mdn	9.72	18.50	11	11	11
Mean	10.60	18.20	10.74	10.61	10.68
S.D.	2.09	3.72	2.54	2.04	2.08

TABLE 21

Correlation of measures of parents' socio-economic status with children's I.Q's at time of first and second test and changes in I.Q. from first to second test.

Ratings	N	MA 1st	MA 2nd	IQ 1st	IQ 2nd	Change in IQ 1st–2nd
Barr Occup.	83	.20 ± .07	.33 ± .07	.17 ± .07	.24 ± .07	.07 ± .07
Whittier	86	.24 ± .07	.18 ± .07	.24 ± .07	.18 ± .07	−.10 ± .07
Educ. Father	85	.18 ± .07	.19 ± .07	.20 ± .07	.25 ± .07	.03 ± .07
Educ. Mother	85	.14 ± .07	.20 ± .07	.23 ± .07	.22 ± .07	−.02 ± .07
Midparent Ed	85	.18 ± .07	.17 ± .07	.24 ± .07	.26 ± .07	.00 ± .07

changes in intelligence quotient after a period of preschool experience offer support to the contention that changes in I.Q. are not environmentally determined or, at least, are not determined by a particular type of environment. It would appear that changes so determined would of necessity bear a significant inverse relationship to measures of the environment from which the child gaining in I.Q. comes to the situation in which I.Q. gains occur. Anderson (3), Goodenough (16), McNemar (27, 28), Simpson (32, 33, 34) and others have requested that Wellman and her colleagues at the University of Iowa "demonstrate the environment" which produces gains in intelligence at the preschool ages. It appears that a first request ought to be "to demonstrate an inverse relationship between environmental status prior to preschool experience and gains in I.Q. resulting from preschool experience." Once this request has been satisfied it ought not to be as difficult to demonstrate differences between the two environments and to set up controlled experiments to measure the effects of various factors, in the preschool environment, which contribute to the gain in I.Q.

In an attempt to evaluate the environmental background from which the child comes to preschool there certainly ought to be included, in a composite score for environment, some evaluation of the school-like experiences of the child prior to beginning preschool. Such experiences, as attending Sunday School, going to summer camp, dancing or music lessons, being enrolled in a play group, etc., would in all likelihood come under the general classification of school-like experiences. A demonstration of this point may be drawn from the present research. Data obtained on 86 of the 91 children were grouped as follows:

Thirty-six children had had "school-like" experiences before entering the public school kindergarten. Of these, 31 had attended Sunday school for periods ranging from six months to 4 years; 4 had attended dancing school for periods ranging from 1 to 2 years; 1 had gone away from home to summer camp for the summer preceding kindergarten entrance.

Fifty of the group were reported by their parents to have had no "school-like" experiences before entrance to kindergarten.

The data below shows comparisons between the "school-like experience" group and the "no school-like experience" group in terms of test re-test I.Q. change.

DATA

	36 children school-like exp.	50 children no school-like exp.
Range	−21 to +23 pts. I.Q.	−13 to +28
Mdn	+3	+7.5
Mean	+3.2	+7.6
S.D.	8.01	8.85

The C.R. for differences obtained between the mean gains of these two groups is 2.40 which approaches significance in view of size of the groups.

A tetrachoric correlation was computed (13, p. 371) which indicates the relationship between having had or not having had school-like experience prior to kindergarten and failure to gain or gaining in I.Q. from test to re-test. A correlation of +.32 was obtained between I.Q. gain and no school-like experience or a positive relationship between school-like experiences and failure to gain in I.Q. from test to re-test after preschool experience. The P.E. of this

r_t calculated as being approximately 1½ times the r for a Pearson Coefficient of correlation of comparable size would be approximately ±.10. An interpretation of this correlation would lead to the conclusion that there is a possible positive relationship between test re-test I.Q. gains after a short period of pre-school experience and no previous school-like experience.

The above data support the suggestion that it ought to be possible to demonstrate a significant inverse relationship between a composite score for the child's socio-economic status and background of experience before preschool and gains in I.Q. score after preschool experience.

Section 7 offers no conclusive evidence of a relationship between the socio-economic and educational status of parents and children's gains in I.Q. after preschool experience. However, evidence is offered to support the hypothesis that such gains in I.Q. are related to the child's experience before entering preschool.

COMPARISONS OF EXTREME GROUPS

In order to make comparisons between extreme groups in terms of gain and failure to gain in I.Q. from test to re-test, statistical data have been computed on 23 children who gained twelve or more I.Q. points each from test to re-test with a mean gain of 16.83, S.D. 4.64; and on 38 children whose second I.Q's were within ±5 points of their initial I.Q. This group made a mean gain of +.79 S.D. 1.50 I.Q. points.

TABLE 22
C.R. of differences between means of gaining and non-gaining group

Measure	C.R.M	Measure	C.R.M
CA's 1st test	.54	Days school attended between tests	2.31
CA's 2nd test	.37	Shyness changes 1–2 tests	.36
Binet MA's 1st test	3.18	Negativism changes 1–2 test	.21
Binet MA's 2nd test	1.52		
Binet I.Q's 1st test	3.27[5]	Social maturity S.Q's	.56
Binet I.Q's 2nd test	2.10		
		Father's occupation Barr Scale	.53
Binet basal age, 1st test	1.50	Home Rating, Whittier Scale	.02
Binet Basal age, 2nd test	.60		
Binet ceiling ages, 1st test	3.32	Mid-Parent education	1.30
Binet ceiling ages, 2nd test	2.07		

According to Table 22 a gaining group of 23 children did not differ significantly from a non-gaining group of 38 children in mean C.A. at the times of either the first or second Binet tests; the gaining group did differ significantly from the non-gaining group in Mean Binet M.A. and I.Q. on the first test. Differences favored the non-gaining group. At the time of the second test the differences favored the gaining group who had a Mean I.Q. 6.21 points higher than the non-gaining group—as con-

[5] The author recognizes these C.R.'s to be statistical artifacts due to regression and resulting from the fact that the groups are selected on basis of gain or failure to gain from test to re-test.

trasted with a Mean I.Q. 10.07 points lower than the non-gaining group on the first test.

The two groups did not differ significantly in Mean basal ages on either test. They did differ significantly in Mean first-test ceiling ages. The difference of 10.06 months is in favor of the non-gaining group. The Mean first test ceiling age of 88.58 months of the non-gaining group carries them well into that area of the 1937 Binet test above the 72 months level in which, according to Table 14, approximately 90% of the items require oral speech of the testee.

The Mean first test ceiling age of 78.52 months of the gaining group does not reach above the first age level, beyond 72 months, in which the test is so heavily weighted with items requiring oral speech. An inspection of Table 14 shows that in both forms of the Binet test the age levels from 72 to 84 months contain 100% of items requiring oral speech of the testee. The data from which Table 22 was computed show that only one child or 4% of the gaining group succeeded in reaching a ceiling age above the 72-84 months level on the first test while 14 or 39% of the non-gaining group reached ceiling ages at 96 months or above. On the second test 18 or 78% of the gaining group reached ceiling ages of 96 months or above. 78% of these children succeeded in going beyond (passing one or more items) an age level made up of oral speech items as contrasted with only 4% going beyond this age level on the first test before preschool experience. 45% of the non-gaining group reached ceilings at 96 months or above on the second test. These data strengthen the belief that gains in I.Q. earned by these children are in part due to the effects of preschool experience on oral speech behavior.

The findings reported in Table 22 with regards to changes in shyness and negativism deal with mean changes of the groups in the direction of being judged to be less shy and negativistic at the time of the second tests.

Section 8 offers additional evidence to support the hypothesis that gains in I.Q. after preschool experience are adjustmental rather than growth in intellectual capacity.

Chapter V

Summary Statement

Ninety-one children with a mean C.A. of 62.02 months were tested before entrance to preschool and re-tested after a mean attendance of 30 three-hour sessions at preschool on the 1937 Revision of the Stanford-Binet Test. Half of the group were initially tested on form L and half on M. Re-tests were on forms alternate to the initial test. Initial and final tests were administered by the same examiner in every case. The mean interval between tests was 1.93 months.

Change in I.Q. from test to re-test has been studied in relation to (1) length of preschool attendance, (2) interval between tests, (3) C.A. at time of initial test, (4) basal and ceiling ages scores on initial and final tests, (5) individual test items passed, failed and refused on initial and final tests, (6) examiners' judgment of rapport, (7) social maturity, and (8) socio-economic and educational status of parents.

Conclusive evidence is presented to show that children do make significant mean gains in I.Q. score as a result of a short period of preschool experience. Since these gains occurred so rapidly it is unlikely that they result from growth in intellectual capacity. It is more probable that they are due to better adjustment of the testee to the testing situation. This means that we have no right to take a single preschool test result of an intelligence test as a measure of intellectual capacity. In this respect we would agree with the findings of Wellman although our interpretation of these findings differs.

No conclusive evidence is offered for a relationship between I.Q. gain and length of preschool attendance or interval between tests. However, trends for a positive relationship are disclosed and there is some indication that the major gain in I.Q. may occur after one month of preschool experience.

Evidence obtained for a relationship between C.A. at the time of the initial tests and I.Q. gain on the final test indicates a trend towards an inverse relationship.

Evidence was obtained to support the hypothesis that I. Q. gains resulting from preschool experience are adjustmental gains rather than growth in intelligence. It is shown that part of the adjustment is improvement in the use of oral speech in the test situation.

No conclusive evidence has been obtained to support the hypothesis that gain in I.Q. is related to improvement in the examiner's rapport with the child on the second test. Slight trends are disclosed in the direction of a relationship between gain in I.Q. and a reduction of shyness and negativism.

Social maturity as measured by the Vineland Social Maturity Scale (11) was not found to be related to I.Q. gains.

Socio-economic and educational status of the parents and home ratings were not found to be significantly related to I.Q. gains. However, evidence is offered for a positive relationship between lack of school-like experience before entrance to preschool and gain in I.Q. after preschool experience. Some evidence is offered which indicates that future research may disclose a significant inverse relationship between a composite score

for socio-economic status, home rating, and the child's background of experience and I.Q. gains resulting from preschool experience.

A group of children who made large gains in I.Q. after preschool experience is compared with a similar group who made no significant I.Q. gains. The groups are not found to differ significantly on any of the measures used but analysis of the individual tests offers even stronger support for the hypothesis that I.Q. gains after preschool experience are related to improvement in the use of oral speech in the test situation.

The final conclusions of this report are:

1. Significant gains in I.Q. score do result from preschool experience when Stanford-Binet Tests are given under standard conditions and when the initial test is given before any preschool experience. Since these gains occur so rapidly it is unlikely that they result from growth in intellectual capacity.

2. I.Q. gains resulting from preschool experience are adjustmental gains rather than growth in intelligence.

BIBLIOGRAPHY

1. Anderson, J. E. The limitations of infant and preschool tests in the measurement of intelligence. *J. Psychol.*, 1939, *8*, 351-379.
2. Anderson, J. E. Prediction of terminal intelligence from infant and pre-school tests. *39th* Year-Book, N.S.S.E., 1940, (I), 385-403.
3. Anderson, L. D. A longitudinal study of the effects of nursery school training on successive intelligence test ratings. *39th Year-Book*, N.S.S.E., 1940, (II), 3-10.
4. Bayley, N. Mental growth in young children. *39th Year-Book*, N.S.S.E., 1940, (II), 11-47.
5. Bayley, N. Factors influencing the growth of intelligence in young children. *39th Year-Book*, N.S.S.E., 1940, (II), 49-79.
6. Bird, G. E. The effect of nursery school attendance upon mental growth of children. *39th Year-Book*, N.S.S.E., 1940, (II), 81-84.
7. Caille, R. K. Resistant behavior of preschool children. *Child. Devel. Monog.*, 1933, No. 11, 142 pp.
8. Cattell, P. The measurement of intelligence of infants and young children. New York: Psychological Corp., 1940.
9. Coffey, H. S., & Wellman, B. L. The role of cultural status in intelligence changes of preschool children. *J. exp. Educ.*, 1936, *5*, 191-202.
10. Dearborn, W. F., & Rothney, J. W. M. Predicting the child's development. Cambridge, Mass.: Sci-Art, 1941.
11. Doll, E. A. The Vineland Social Maturity Scale. *Vineland Training School Series*, 1936, No. 3.
12. Frandsen, A., & Barlow, F. P. Influence of the nursery school on mental growth. *39th Year-Book*, N.S.S.E., 1940, (II), 143-147.
13. Garrett, H. E. Statistics in psychology and education. New York: Longmans Green, 1939.
14. Goodenough, F. L. Look to the evidence: a critique of recent experiments on raising the I.Q. *Educ. Meth.*, 1939, *19*, 73-79.
15. Goodenough, F. L. Can we influence mental growth. A critique of recent experiments. *Educ. Rec.*, 1940, *21*, suppl. 13, 120-143.
16. Goodenough, F. L. New evidence on environmental influence on intelligence. *39th Year-Book*, N.S.S.E., 1940, (I), 307-356.
17. Goodenough, F. L. Measurement of intelligence by drawings. New York: World Book Company, 1926.
18. Goodenough, F. L., & Maurer, K. M. The mental development of nursery school children compared with that of non-nursery school children. *39th Year-Book*, N.S.S.E., 1940, (II), 161-178.
19. Goodenough, F. L., & Maurer, K. M. The relative potency of the nursery school and the statistical laboratory in boosting the I.Q. *J. educ. Psychol.*, 1940, *31*, 541-549.
20. Hallowell, D. K. Stability of mental test ratings for pre-school children. *J. genet. Psychol.* 1932, *40*, 406-421.
21. Hollingworth, L. S. The significance of deviates. *39th Year-Book*, N.S.S.E., 1940, (I), 43-66.
22. Honzik, M. P. The constancy of mental test performance during the pre-school period. *J. genet. Psychol.*, 1938, *52*, 285-302.
23. Honzik, M. P. Age changes in the relationship between certain environmental variables and children's intelligence. *39th Year-Book* N.S.S.E. 1940, (II), 185-205.
24. Jersild, A. T., & Fite, M. D. The influence of nursery school experience on children's social adjustments. *Child Devel. Monog.* 1939, No. 25.
25. Lamson, E. A. A follow-up study of nursery school children. *39th Year-Book* N.S.S.E., 1940, (II), 231-236.
26. Levy, D. M., & Tulchin, S. H. The resistant behavior of infants and children, *J. exp. Psychol.* 1925, *8*, 223.
27. McNemar, Q. A critical examination of the University of Iowa studies of environmental influences upon the I.Q. *Psychol. Bull.*, 1940, *37*, 63-92.
28. McNemar, Q. More on the Iowa studies, *J. Psychol.*, 1940, *10*, 237-240.
29. Olson, W. C., & Hughes, B. O. Subsequent growth of children with and without nursery school experience. *39th Year-Book* N.S.S.E., 1940, (II), 237-244.
30. Reynolds, M. M. Negativism of pre-school children. *Teach. Coll. Contr. Educ.*, 1928, No. 288.
31. Rust, M. M. The effects of resistance on intelligence test scores of young children. *Child Devel. Monog.* 1931, No. 6. 80 pp.
32. Simpson, B. R. The wandering I.Q. Is it time for it to settle down? *J. Psychol.*, 1939, *7*, 351-367.
33. Simpson, B. R. The wandering I.Q.; fact or statistical illusion? *Sch. and Soc.*, 1939, *50*, 20-23.
34. Simpson, B. R. The wandering I.Q.; a continuation. *J. Psychol.* 1940, *9*, 31-48.
35. Starkweather, E., & Roberts, K. I.Q. changes occurring during nursery school attend-

ance at the Merrill-Palmer School. *39th Year-Book* N.S.S.E., 1940, (II), 315-335.

36. TERMAN, L. M. *Genetic studies of genius,* (I) Stanford Univ. Press, 1925.
37. TERMAN, L. M., & MERRILL, M. A. *Measuring intelligence.* New York: Houghton Mifflin, 1937.
38. UPDEGRAFF, R. The determination of a reliable I.Q. for the young child. *J. genet. Psychol.,* 1932, *41,* 152-166.
39. VOAS, W. H. Does attendance at the Winnetka Nursery School tend to raise the I.Q.? *39th Year-Book* N.S.S.E., 1940, (II), 363-376.
40. WELLMAN, B. L. Some new bases for interpretation of the I.Q. *J. genet. Psychol.* 1932, *41,* 116-126.
41. WELLMAN, B. L. The effect of preschool attendance upon the I.Q., *J. exp. Educ.,* 1932-33, *1,* 48-69.
42. WELLMAN, B. L. Growth in intelligence under differing school environments, *J. exp. Educ.,* 1934-35, *3,* 59-83.
43. WELLMAN, B. L. Mental growth from preschool to college. *J. Exp. Educ.* 1936-37, *6,* 127-138.
44. WELLMAN, B. L. Our changing concept of intelligence. *J. consulting Psychol.,* 1938, *2,* 97-107.
45. WELLMAN, B. L. The intelligence of preschool children as measured by the Merrill-Palmer Scale of performance tests., *Univ. Ia. Stud. Child Welf.,* 1938, *15,* No. 3, 150 pp.
46. WELLMAN, *B.* L. Changing concept of the I.Q. *J. Home Economics,* 1939, *31,* 77-80.
47. WELLMAN, B. L. The I.Q.; a reply. *J. Psychol.,* 1939, *8,* 143-155.
48. WELLMAN, B. L. Guiding mental development, *Childhood Educ.,* 1938-39, *15,* 108-112.
49. WELLMAN, B. L. The fickle I.Q. *Sigma Xi Quart.* 1940, 51-60.
50. WELLMAN, B. L. The meaning environment. *39th Year-Book* N.S.S.E., 1940, (II), 21-36.
51. WELLMAN, B. L. Studies on effect of schooling. *39th Year-Book,* N.S.S.E., 1940, (II), 377-399.
52. WILLIAMS, J. H. Whittier scale for grading home conditions, *J. juv. Res.,* 1916, *1,* 271-286.